A Model of the Switched Telephone Network for Data Communications

By E. O. ELLIOTT

(Manuscript received August 20, 1964)

The error statistics from data-transmission field tests on the telephone network may be compactly represented by about one dozen parameters. These relate to a model of the telephone network in which there are three distinct channels. The errors in binary data on each channel are produced by a renewal process in which a bit-error is a renewal event. The mixture of three such channels allows a close fit to the error statistics for a large range of block lengths. It is not implied that the errors on the telephone network are actually produced by such processes, but merely that they may be conveniently and compactly represented by them.

Use of this model simplifies the analysis of error-control systems and the determination of error rates for error-control codes. In this paper the model is applied to study the effect that interleaving (time division multiplexing) has on the effectiveness of error-correcting codes.

I. INTRODUCTION

In the study of errors on data communication channels, several mathematical models of the error process bave been proposed. Gilbert[1] proposed a two-state Markov process, and Berger and Mandelbrot[2] have employed a Pareto distribution to fit data collected from the German telephone network. Sussman[3] has applied a Pareto fitting to part of the Alexander-Gryb-Nast data.[4] Common to the models of Gilbert and Berger-Mandelbrot is the assumption that the error process is of the renewal type wherein the state of a bit-error is the renewal event whose occurrence frees the process from dependence upon past history and starts it anew. In such models the distributions of lengths of error-free intervals (gaps) determine the processes, because the lengths of the gaps before and after an error are independently distributed. One may calculate from this distribution tbe probabilities $P(m,n)$ that m bit-errors occur in a block of n consecutive bits. Tbese probabilities are useful

in the analysis of error-control methods for data communications systems.[5]

On the telephone network, as exemplified by the Alexander-Gryh-Nast (AGN) data, the $P(m,n)$ for individual calls are, as noted in Ref. 5, quite diverse in nature. This suggests that there are several error processes involved. In fact, the combined AGN data cannot be described by just one error process of the renewal type but require the mixture of several such processes for a satisfactory description. The mixture of a collection of processes is determined by a corresponding set of channels on which the separate error processes act and specification of the probabilities of being assigned the various channels when placing a call (the assignment, and hence the error process acting, is fixed for the duration of the call).

The purpose of the present paper is to show that in this way the use of three distributions for gap lengths can yield satisfactory approximations to the probabilities $P(m,n)$ for the telephone network. This cannot be accomplished with just one or two distributions, and conceivably four or more might be required in some cases. The $P(m,n)$ for renewal processes depend heavily on the first few values of the gap-length distribution. Because of this there is some choice of distributions for satisfactory models of the telephone network. Noting how various gap-length distributions affect the $P(m,n)$ distributions, we select three gap-length distributions with appropriate weightings to represent the combined switched telephone network. This selection gives excellent agreement to $P(m,n)$ over a wide range of block lengths n. To exemplify further this means of representing the telephone network we apply it also to the Townsend-Watts (TW) data given in Ref. 6.

In the present paper we do not attempt to optimize the degree to which the models represent the telephone network. Rather, we attempt to demonstrate the feasibility of such representations and note that a better accuracy of fit would hardly affect the applications suggested. Also, no attempt is made to associate the parameters with particular causes, such as type of exchange or calling distance, etc. These goals would be the object for future work. We do suggest, however, that dropouts (momentary open-line conditions) and the test words used in field tests are at least partially the cause for the hump in the $P(m,n)$ curves at m near $n/2$.

II. RENEWAL-TYPE ERROR PROCESSES

For a renewal error process, the lengths of successive gaps are independent and distributed according to a common distribution. Let $p(k)$

be the probability that a gap length is $k - 1$, i.e., $p(k) = \Pr(0^{k-1}1 \mid 1)$ where 1 denotes an error-bit, 0 a correct bit and 0^i denotes i consecutive 0's.

Let

$$P(k) = \sum_{m=k}^{\infty} p(m)$$

so that $P(k)$ is the probability that at least $k - 1$ 0's follow a given error [i.e., $P(k) = \Pr(0^{k-1} \mid 1)$]. Now, if p_1 is the unconditional probability of a bit-error, then $p_1 P(k) = \Pr(1)\Pr(0^{k-1} \mid 1)$ is the probability of 10^{k-1}. But, because of the independence among gap-lengths of a renewal process, order is irrelevant and it is clear that the events 10^{k-1} and $0^{k-1}1$ are equiprobable. Hence, $p_1 P(k) = \Pr(0^{k-1}1)$. To obtain the value of p_1, note that $p_1 = 1/\bar{k}$ where $\bar{k}$, the average distance to the next error, is equal to

$$\sum_{k=1}^{\infty} kp(k).$$

The probabilities of individual error patterns of a renewal process are easily calculated (but we do not make use of these here). For example, consider a block ζ of n consecutive bits which contains m bit-errors and, as in Ref. 5, p. 1985, let a be the number of 0's before the first 1 in ζ, c the number of 0's following the last 1 in ζ and b_i ($i = 1, \cdots, m - 1$) be the number of 0's between consecutive 1's in ζ. Then, the probability of ζ's occurrence is given by

$$\Pr(\zeta) = p_1 P(a + 1) \left\{ \prod_{i=1}^{m-1} p(b_i + 1) \right\} P(c + 1).$$

Calculations of the above sort may be of use in evaluating both error-correcting and error-detecting codes on renewal-type channels. For more general, but approximate, applications, the probabilities $P(m,n)$ that m bit-errors occur in a block of length n are of use. To calculate these we may use recurrence relations or generating functions as follows.

First, let $R(m,n)$ be the probability that $m - 1$ errors occur in the next $n - 1$ bits following an error. Thus, $R(1,n) = P(n)$ for $n \geqq 1$, and

$$R(m,n) = \sum_{k=1}^{n-m+1} p(k) R(m - 1, n - k)$$

for $2 \leqq m \leqq n$.

Now,

$$P(m,n) = \sum_{k=1}^{n-m+1} p_1 P(k) R(m,n-k+1)$$

whenever $1 \leqq m \leqq n$.

Computer programs for computing $P(m,n)$ from the above recurrence relations have been written and used in obtaining the data presented later in this paper.

An alternate approach to the above relation is through generating functions. If we let $g(z)$ and $G(z)$ be the generating functions associated with $p(k)$ and $P(k+1)$ respectively

$$\left(\text{i.e., } g(z) = \sum_{k=1}^{\infty} p(k)z^k \quad \text{and} \quad G(z) = \sum_{k=0}^{\infty} P(k+1)z^k\right)$$

then, from Ref. 7, p. 249, we have

$$G(z) = \frac{1-g(z)}{1-z}.$$

Letting

$$H_m(z) = \sum_{n=m}^{\infty} P(m,n)z^n$$

we obtain that

$$H_m(z) = p_1 z G(z) g(z)^{m-1} G(z)$$

considering that m errors involve $m-1$ gaps in a total number of bits adding up to n and that the generating function for a convolution of variables is the product of their associated generating functions. [$p_1 z G(z)$ is the generating function for the probabilities of the events $0^{k-1}1$, $g(z)$ is that of $0^{k-1}1 \mid 1$ and $G(z)$ is that of $0^k \mid 1$.]

Thus, we obtain

$$H_m(z) = p_1 z \left\{\frac{1-g(z)}{1-z}\right\}^2 g(z)^{m-1}.$$

Calculation of $P(m,n)$ from this generating function is rather inconvenient. The recurrence equations are generally preferred in practice.

III. A REPRESENTATION OF THE TELEPHONE NETWORK

In both of the data-transmission field-test programs on the telephone network, data calls were placed on a variety of circuits and bit-errors were recorded. In Refs. 5 and 6 the composite effects of these are represented by the $P(m,n)$ probabilities. Fig. 1 shows $P(m,31)$ for these two

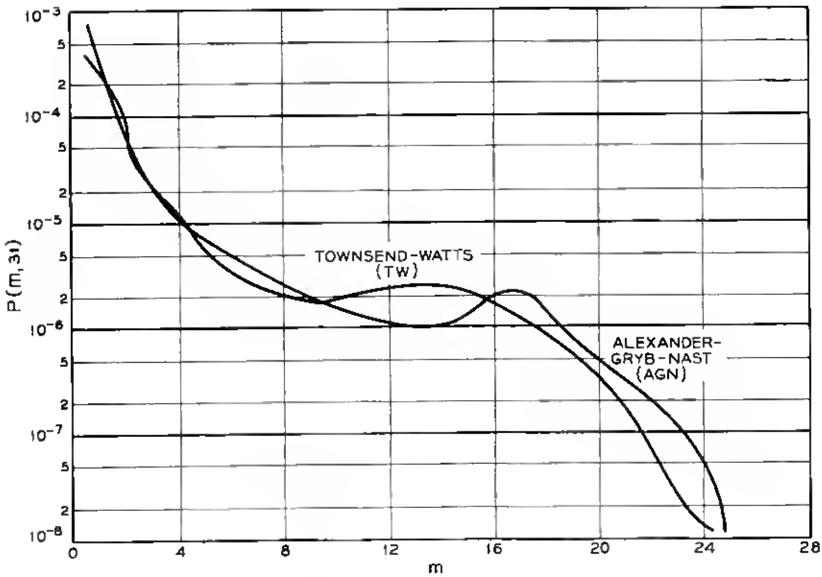

Fig. 1 — $P(m,31)$ for field test data.

field tests and is typical of $P(m,n)$ for other intermediate block lengths n. These curves appear to have three separate segments: an initial segment with a steep slope, an intermediate segment with a smaller slope, and a terminal hump and tail. Using the recurrence equations of the previous section, the $P(m,31)$ curves for the three gap-length distributions (determined by trial and error) given in Table I were calculated and are displayed in Fig. 2. The unconditional bit-error rate for each curve is taken to be that of the AGN data. In so doing we essentially assume that the tails of the gap-length distributions are appropriately tailored. The tails of these distributions, of course, do not influence $P(m,n)$ when n is not too large. Each of these three curves is more-or-less parallel to the respective first, second or third segment of the $P(m,31)$ curve for the AGN data. They are to be added together, after multiplication by suitable weighting factors, to produce our approximation to the AGN curve. By trial and error, we find that weighting factors of 50, 25, and 25 per cent respectively give the close fit which is shown in Fig. 3 for block length 31 and again in Figs. 4–7 for some other block lengths. This trial and error procedure of finding curves of the right shape and then appropriate weighting factors represents a simple attempt to approximate the $P(m,n)$ curves of the AGN data

TABLE I — GAP-LENGTH DISTRIBUTIONS $p(k)$ FOR AGN MODEL

$k =$	1	2	3	4	5	>5
Initial segment	0.12	0.06	0.0	0.0	0.0	≈0
Intermediate segment	0.40	0.20	0.10	0.05	0.0	≈0
Hump	0.56	0.24	0.06	0.0	0.12	≈0

over a wide range of block lengths n. It would be desirable to use analytic methods instead of trial and error in obtaining such approximations, but the number of parameters involved is large and the analytic expressions for $P(m,n)$ are cumbersome. Undoubtedly, appropriate programming techniques can be developed to improve upon our trial and error method.

The gap-length distributions and weighting factors given in Table II furnish an approximate model for the TW data. Figs. 8–10 compare the original and the model at block lengths 10, 31, and 63.

IV. CHOICE AND INTERPRETATION OF COMPONENT DISTRIBUTIONS

Suppose the field test data we wish to represent by a mixture of different renewal processes have an unconditional bit-error rate p_1 and suppose that gap-length distributions $p^i(k)$ (the superscript is not an

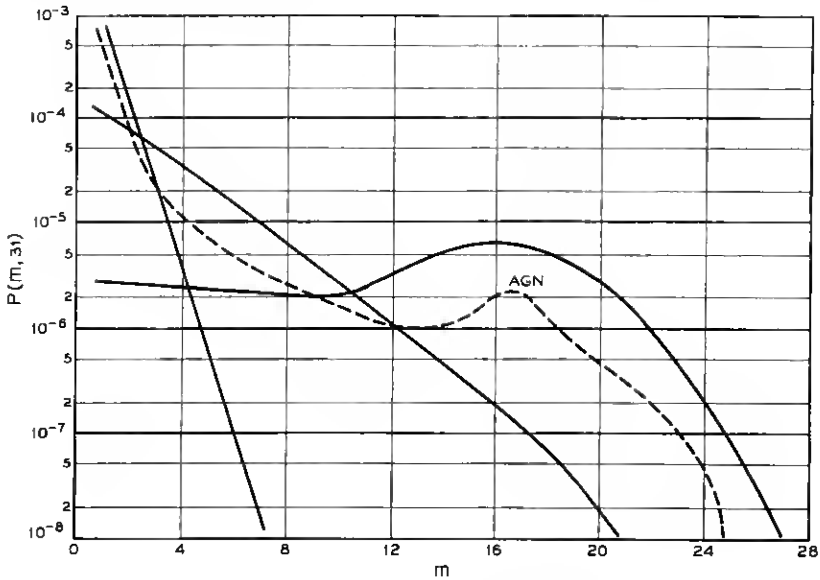

Fig. 2 — $P(m,31)$ for components of model.

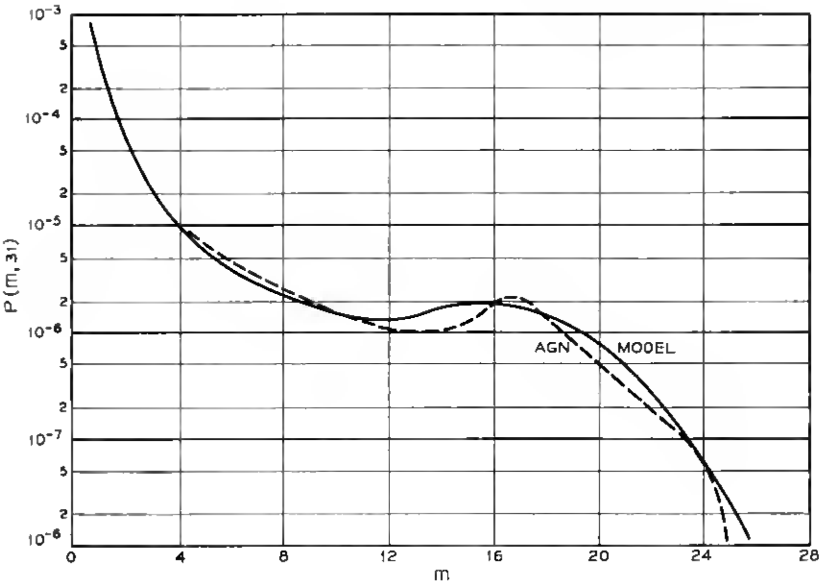

Fig. 3 — $P(m,31)$ for model of AGN data.

exponent) and weighting factors λ_i ($i = 1, \cdots, J$) have been decided upon. (Determination of $p^i(k)$ and λ_i will be discussed subsequently.) Then, let $P^i(m,n)$ be calculated for each distribution $p^i(k)$ using the prescribed error rate p_1. For the model we then take

$$P(m,n) = \sum_{i=1}^{J} \lambda_i P^i(m,n) .$$

Use of the common value p_1 in computing $P^i(m,n)$ does not imply that each distribution $p^i(k)$ has this bit-error rate, but is just a device to assure that the model has the same unconditional bit-error rate as the field-test data. In fact, since $p^i(k)$ is generally specified for only a few small values of k, we assume that $p^i(k)$ for larger values of k is distributed so that its real bit-error rate $p_{1,i}$ may be different from p_1, and it is not used explicitly in our model. Incidentally, $\lambda_i p_1$ represents the portion of the total error rate attributable to the channels with gap-length distribution $p^i(k)$ and we could say that the percentage of channels of this type is λ'_i where $\lambda'_i p_{1,i} = \lambda_i p_1$.

To choose candidate distributions $p^i(k)$ we must first observe some principles, pertaining separately to the segments $i = 1,2,3$, which are noted in the next few paragraphs. We begin by examining the simplest

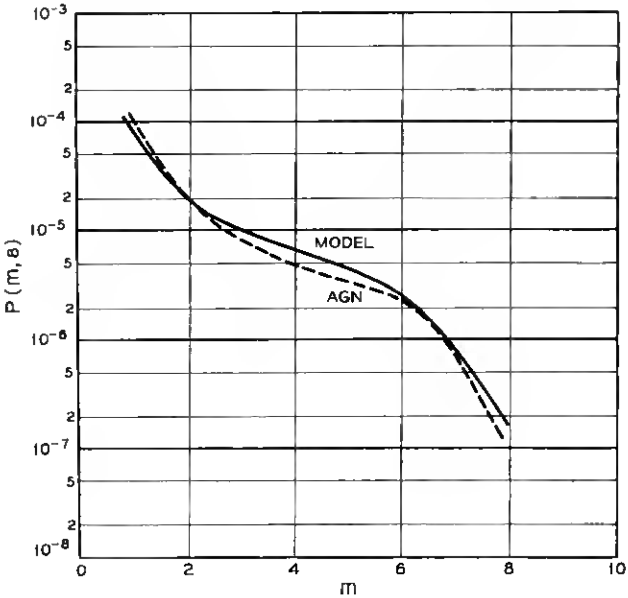

Fig. 4 — $P(m,8)$ for model of AGN data.

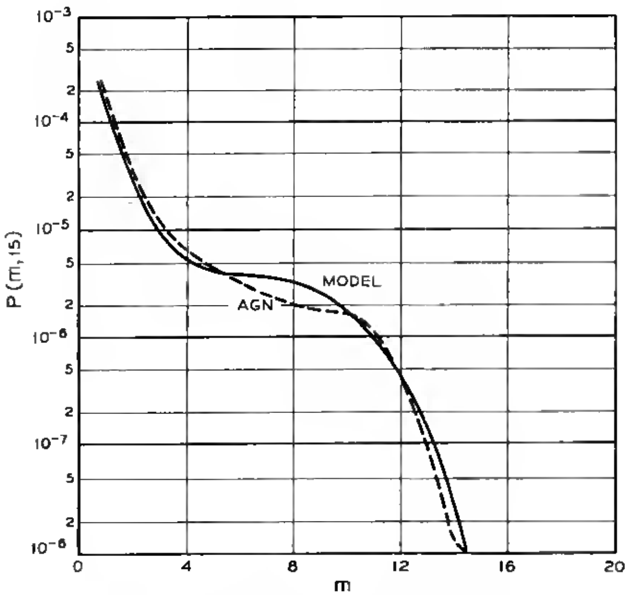

Fig. 5 — $P(m,15)$ for model of AGN data.

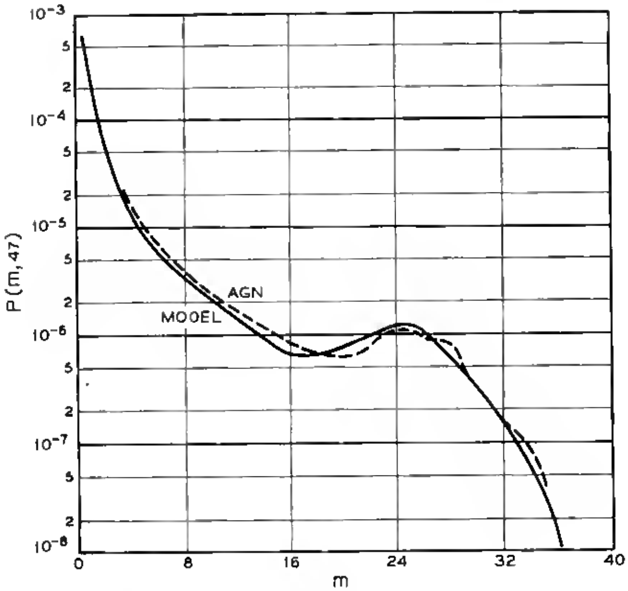

Fig. 6 — $P(m,47)$ for model of AGN data.

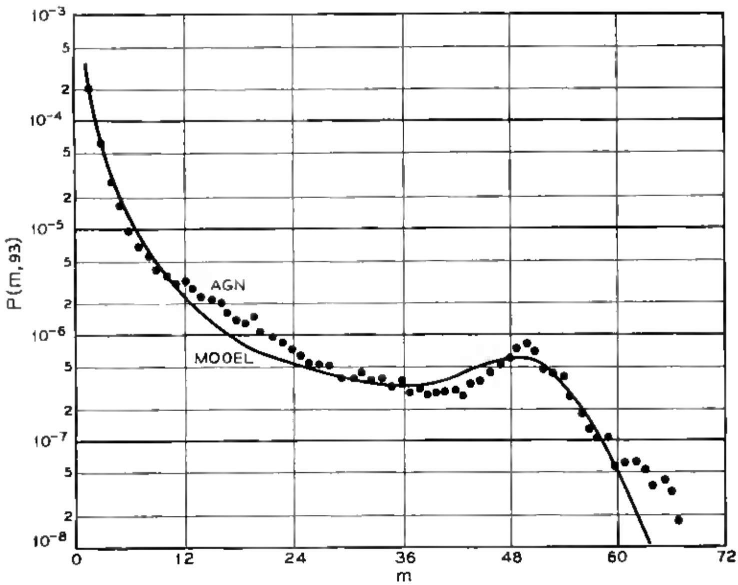

Fig. 7 — $P(m,93)$ for model of AGN data.

TABLE II — GAP-LENGTH DISTRIBUTIONS $p(k)$ AND WEIGHTING FACTORS FOR TW MODEL

Weighting Factor	λ	k					
		1	2	3	4	5	>5
Initial segment	57%	0.20	0.10	0.0	0.0	0.0	≈0
Intermediate segment	18%	0.35	0.25	0.15	0.05	0.0	≈0
Hump	25%	0.45	0.25	0.15	0.10	0.03	≈0

case analytically, namely that of a gap-length distribution $p(k)$ such that $p(1) = \alpha$ and $p(k) \approx 0$ for $k > 1$. Then the generating function $g(z)$ for $p(k)$ is essentially αz and

$$H_m(z) = p_1 z \left[\frac{1 - \alpha z}{1 - z}\right]^2 (\alpha z)^{m-1}.$$

Determining the coefficient of z^n in the above [using $(1 - z)^{-2} = 1 + 2z + 3z^2 + 4z^3 + \cdots$] we obtain

$$P(m,n) = p_1 \alpha^{m-1}[(n - m + 1) - 2\alpha(n - m) + \alpha^2(n - m - 1)]$$

which in a logarithmic plot as a function of m is almost a straight line.

A distribution of this type with α small, or a slight modification

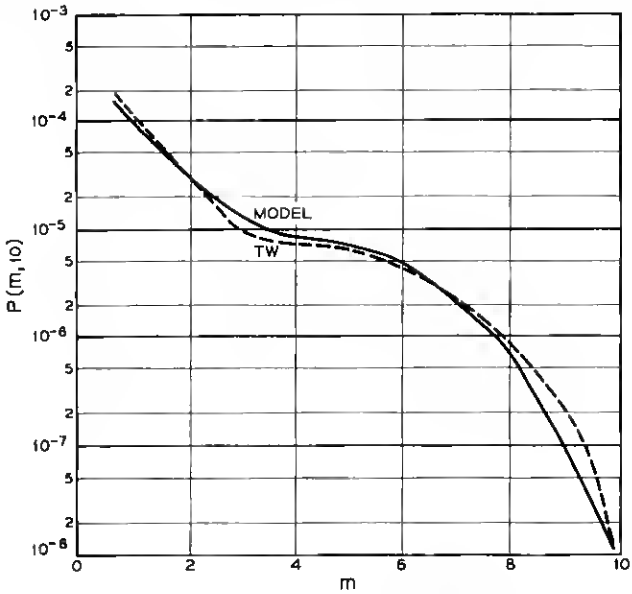

Fig. 8 — $P(m,10)$ for model of TW data.

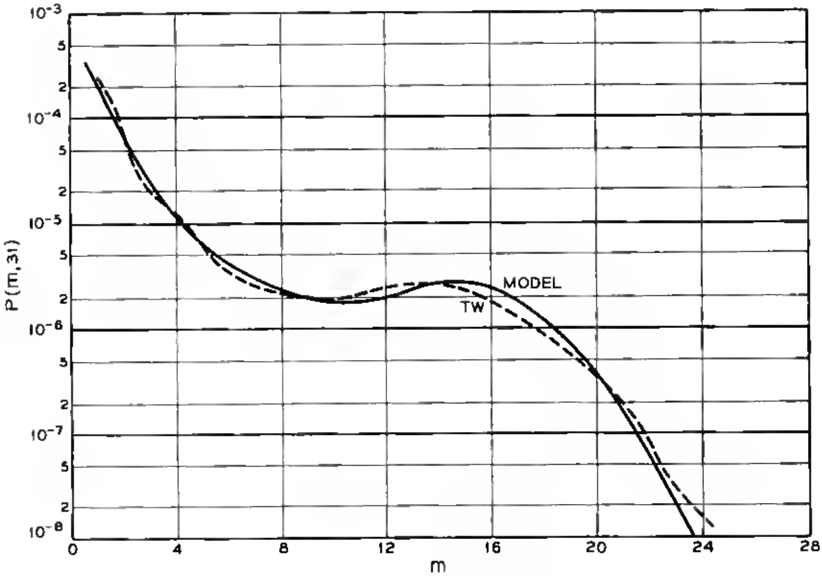

Fig. 9 — $P(m,31)$ for model of TW data.

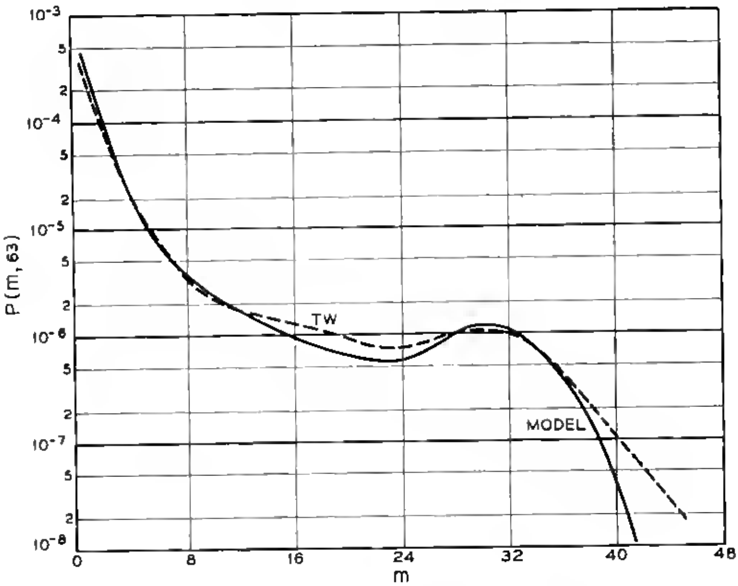

Fig. 10 — $P(m,63)$ for model of TW data.

thereof, may be useful in matching the initial segment of field test data as in the foregoing examples. In general, the initial segment of a $P(m,n)$ curve which has the same general shape as our two field test examples is determined by a gap-length distribution $p(k)$ having 5 to 40 per cent of the probability spread over the first three or four values of k. It is reasonable to assume also that $p(k)$ is monotonically decreasing and that $p(k) \approx 0$ for k much larger than 4.

For the intermediate segment of such a curve, the gap-length distribution may contain between 70 and 80 per cent of the probability in the first four or five terms. A definite hump begins to appear in the $P(m,n)$ curve when much over 90 per cent is contained in the same range. Amounts of 95 to 99 per cent produce humps as extreme as we note in our examples.

The exact way in which $p(k)$ is distributed over these first few terms can influence the shape of the $P(m,n)$ curve considerably. In general, large values of $p(1)$ cause the $P(m,n)$ to remain large for a more extended range of values of m. Beyond these few general remarks, the process of fitting remains a matter of trial and error. First we find gap-length distributions which give rise to $P(m,n)$ curves which approximate in shape the various segments of the curve we are trying to match. Then weighting factors are chosen so that when the various components are so weighted and added together, they yield numerical agreement with the desired curve.

Berger and Mandelbrot[2] and Sussman[3] have claimed that the error processes on the telephone network are indeed of the renewal type and they have some data[2] to support this view. We have not investigated this matter, but we do note that for the AGN data the composite gap-length distribution given in Table III is approximated reasonably well by that for our model. Our model, however, has $p(k) \approx 0$ for $k > 5$, whereas for the AGN data $p(k)$ 0 $\not\approx$ for $k > 5$. These observations neither confirm or deny the existence of renewal processes on the telephone network. The accuracy with which our model approximates the field data is, however, indirect support for the notion that the error processes are at least not widely different from renewal processes.

The humps in the $P(m,n)$ curves at $m \approx n/2$ are rather remarkable. We have suspected that they are at least partially due to drop-outs and to the nature of the error-recording procedures in the field tests. For example, in the AGN tests during a drop-out, the error pattern recorded would coincide with the test word (Ref. 8, p. 1402) and would appear as repetitions of

$$111101111001011000010000110110.$$

TABLE III — GAP-LENGTH DISTRIBUTIONS FOR AGN DATA AND MODEL

$k =$	1	2	3	4	5
AGN data	0.24	0.09	0.05	0.02	0.03
Model	0.20	0.14	0.04	0.01	0.03

When this pattern is present, the gap-length distribution induced would be that given in Table IV. This is essentially the distribution we have used to produce the hump in our model for the AGN data (cf. Table I). We found that it gave a slightly better fit than some other distributions we tried, but other distributions did yield rather good fittings. This is not sufficient evidence to conclude that the hump is entirely due to drop-outs but it does support the hypothesis that it is at least partially caused by them.

One other tentative interpretation of our models for the AGN and TW data lies in the gap-length distributions for their initial segments. The term $p(1)$ for the TW data is almost twice as large as that for the AGN data, which means that double errors would be more prevalent in the TW data. This is consistent with the occurrence of dihit errors for the four-phase data set employed in the TW tests.[6]

So many gap-length distributions seem to give reasonable approximations to the intermediate segment of the $P(m,n)$ curves of these field tests that it is difficult to ascribe much significance to them. They do, however, represent instances where errors have high probabilities of following other errors, thereby producing bursts. Very short drop-outs would offer one explanation, but there are no doubt others.

The accuracy of our approximations is certainly sufficient for many purposes, and in particular for the estimation of error rates for codes by the methods given in Ref. 5. An advantage of the models which is somewhat independent of accuracy is the following. When the block length n is very large there are two defects in the $P(m,n)$ values obtained directly from the field-test data. First, the plot of $P(m,n)$ becomes erratic due to the small sample size afforded by the field-test data, and, second, the computer processing time to obtain $P(m,n)$ becomes excessive. On the other hand, $P(m,n)$ computed from the model yields a

TABLE IV — GAP-LENGTH DISTRIBUTION FOR TEST WORD

k	1	2	3	4	5	>5
$p(k)$	0.563	0.250	0.062	0.0	0.125	0.0

smooth curve with very little time required in the computations. This appears to be a significant advantage of representing the telephone network by way of mathematical models.

V. ERROR RATES ON THE TELEPHONE NETWORK

The use of $P(m,n)$ in estimating error rates for error-detecting codes is described in Ref. 5. It has also been used for some error-correcting codes in Ref. 9. For error-correcting codes, the probability P_e of incorrect decoding may be conveniently and usefully bounded in some cases. For example, if the code is capable of correcting all $(m - 1)$-fold (or fewer) errors, then certainly $P_e \leqq P(\geqq m,n)$ where

$$P(\geqq m,n) = \sum_{i=m}^{n} P(i,n)$$

and n is the code's block length.

In Ref. 5 we find another use for the probabilities $P(\geqq m,n)$. For an error-detecting code used with retransmissions for error correction, let P_u be the probability of an undetected error and P_r be the probability of retransmission. Then,

$$P_u \approx 2^{-c}P(\geqq d,n)$$

and

$$P_r \approx P(\geqq 1,n)$$

where d is the minimum distance of the code and c is the number of check bits in each code word. The above approximation for P_u is best if c is small, n is moderately large, and the check bits are not too trivial. It may be used in some other cases but with special caution if either extremely low error rates are desired or n is quite large.

Using the model for the AGN data, $P(\geqq m,n)$ vs n has been computed and displayed in Fig. 11 for $m = 1, \cdots, 10$. Fig. 11 is useful for the kind of estimates indicated above and for other considerations in error-control systems (e.g., synchronization).

VI. THE EFFECT OF INTERLEAVING ON ERROR RATES

Time division multiplexing (interleaving) has often been considered as a means of enhancing the error-control effectiveness of error-correcting codes. Its effect on the error statistics of a Gilbert burst-noise channel were noted in Ref. 5, p. 1987.

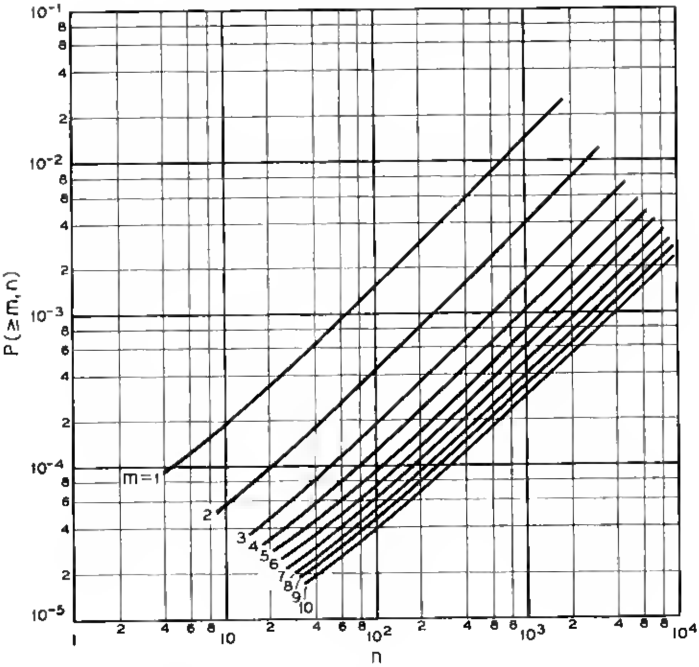

Fig. 11 — $P(\geqq m,n)$ for $m = 1, \cdots, 10$.

With interleaving, the blocked bits from the data source are rearranged (by some delay and storage device) and put onto the line so that of two originally adjacent bits in a block, the second is the tth bit on line following the first. The number t will be called the "interleaving constant." When $t = 1$ there is no interleaving. We let n be the block length, and for error-rate considerations we are concerned with the probabilities $P_t(m,n)$ that m bit errors occur among the n bits of an interleaved data block. For $P_1(m,n)$ we write simply $P(m,n)$ as before. Thus, $P_t(m,n)$ is the probability of m errors among n bits which are equally spaced with $t - 1$ other bits between each two.

To obtain m errors among n bits spaced $t - 1$ bits apart, we must have r errors in the total block of tn bits where $m \leqq r \leqq (t - 1)\, n + m$. Given a total of r errors in tn bits, there are $\binom{r}{m}\binom{tn - r}{n - m}$ collections of n bits (from the total) containing m errors. There are, however, $\binom{tn}{n}$

possible collections of n bits. The probability that n bits selected from the tn bits contain m errors is bence

$$\binom{r}{m}\binom{tn-r}{n-m} \div \binom{tn}{n}$$

and, assuming that a sample of n regularly spaced bits is similar to a random sample of n bits, we may conclude that

$$P_t(m,n) \approx \sum_{r=m}^{(t-1)n+m} \frac{\binom{r}{m}\binom{tn-r}{n-m}}{\binom{tn}{n}} P(r,tn). \tag{1}$$

This formula could be used to approximate $P_t(m,n)$ from the $P(r,tn)$ values, except that $N = tn$ may be quite large and the computation of $P(r,N)$ then becomes infeasible even with the efficient programs used in connection with our mathematical model of the telephone network. Let us abandon this approach.

Now, $P(r,N)$ is expressed by

$$P(r,N) = \sum_{i=1}^{3} \lambda_i P^i(r,N)$$

where $P^i(r,N)$ is associated with the renewal channel determined by the gap-length distribution $p^i(j)$. If we interleave on one of these renewal cbannels, we effectively have anotber renewal cbannel determined by a modified gap-lengtb distribution $p_t^i(j)$. Expressions for this will be given below. Thus,

$$P_t(m,n) = \sum_{i=1}^{3} \lambda_i P_t^i(m,n) \tag{2}$$

where $P_t^i(m,n)$ is obtained from $p_t^i(j)$ in the same way that $p^i(r,N)$ is obtained from $p^i(j)$ and involves only modest computations.

Using (2) we are then capable of computing the $P_t(m,n)$ values for the switched telephone network that are displayed and discussed below.

For a renewal channel the gap-length distribution $p(j)$ specifies the probability $Pr(0^{j-1}1 \mid 1)$ that, given a bit error, the next error is the jth following bit. (The superscript i is dropped from $p(j)$ here, since the three cases $i = 1$ to 3 are treated the same.) Let a be the autocorrelation function for the process so that $a(j) = \Pr(\alpha 1 \mid 1)$ where α is any binary word of length $j - 1$. Then

$$a(1) = p(1)$$

and

$$a(j) = p(j) + \sum_{s=1}^{j-1} p(s)a(j - s) \tag{3}$$

for $j > 1$.

Now, if we interleave with a constant t, error bits remain renewal events and $p_t(j)$ is the probability

$$\Pr\,(\alpha_1 0 \alpha_2 0 \cdots \alpha_{j-1} 0 \alpha_j 1 \mid 1)$$

that looking at every tth bit there are $j - 1$ correct bits preceding the next error ($\alpha_1, \alpha_2, \cdots, \alpha_j$ are arbitrary binary sequences of length $t - 1$).

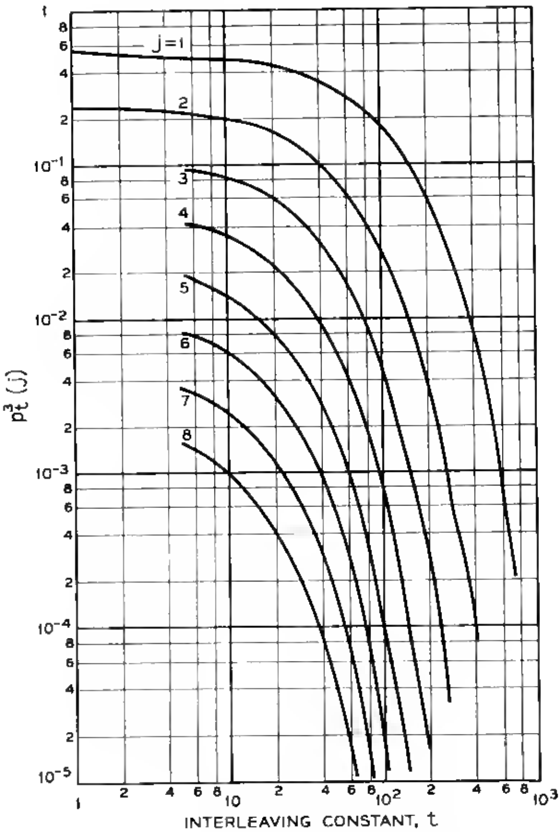

Fig. 12 — Gap-length distribution $p_t^3(j)$ vs t.

Thus,

$$p_t(1) = a(t)$$

and

$$p_t(j) = a(tj) - \sum_{s=1}^{j-1} a(ts)p_t(j-s) \tag{4}$$

for $j > 1$.

This last equation follows from the fact that the first term on the right is the probability that the tjth bit following an error is also in error, and the second term is the total probability that at least one of the tth, $(2t)$th, $\cdots$, $(j-1)t$th bits is in error. Their difference then is $p_t(j)$.

We reintroduce the superscript i $(=1,2,3)$ on $p^i(j)$ to denote respectively the initial, intermediate and hump gap-length distributions given in Table I. The corresponding autocorrelation functions determined by (3) turn out to be well approximated (at multiples of 25) by the following formulae.

$$a^1(j\cdot 25) = 1.43 \times 10^{-13}\{2.295 \times 10^{-13}\}^{j-1}$$
$$a^2(j\cdot 25) = 0.990\{1.863 \times 10^{-2}\}^{j-1}$$
$$a^3(j\cdot 25) = 0.407\{0.763\}^{j-1}.$$

Table V — Gap-Length Distributions $p_t^i(j)$

t	j	$i = 1$	2	3
25	1	0	0.01	0.41
	2	0	0	0.14
	3	0	0	0.05
	4	0	0	0.02
	5	0	0	0.01
50	1	0	0.0002	0.31
	2	0	0	0.08
	3	0	0	0.02
	4	0	0	0.01
100	1	0	0	0.18
	2	0	0	0.03
	3	0	0	0.005
200	1	0	0	0.06
	2	0	0	0.004
	3	0	0	0.0002
400	1	0	0	0.008
	2	0	0	0.0001

Using (4), these give rise to the following gap-length distributions $p_t^i(j)$ (for $t = 25$, 50, 100, 200 and 400) shown in Table V (therein 0 indicates a number less than 10^{-4}).

To illustrate further the dependence on t we have plotted $p_t^3(j)$ ($j = 1$ to 8) in Fig. 12.

Using the gap-length distributions in Table V, we have computed $P_t(n,m)$ for $n = 15$, 31, 93 to obtain $P_t(\geqq m,n)$. These are summarized in Figs. 13 and 14, which give $P_t(\geqq m,31)$ vs t for $m = 1$ to 7 and $P_t(\geqq 3,n)$ vs n for the values of t in Table V.

The correction of all single and double errors is generally considered to be a reasonable procedure for capable codes of almost any block length. For such codes, the probability P_e of a decoding error cannot

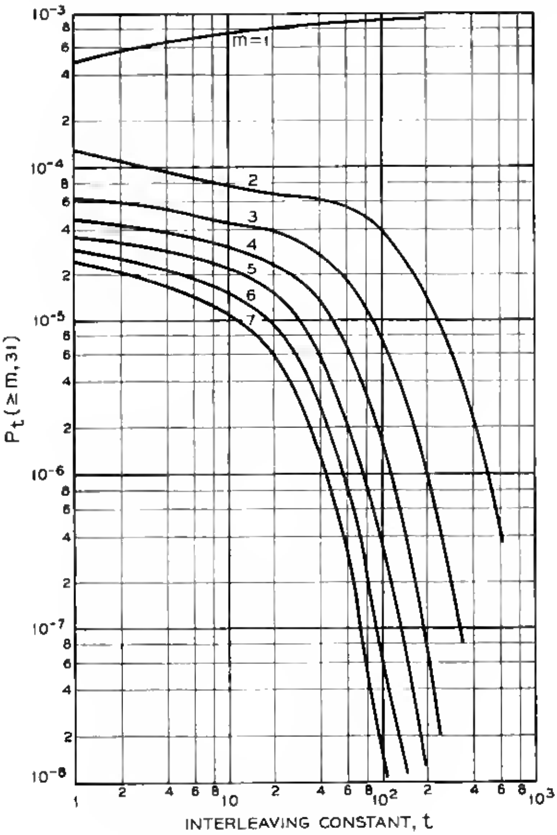

Fig. 13 — $P_t(\geqq m,31)$ vs t for $m = 1, \cdots, 7$.

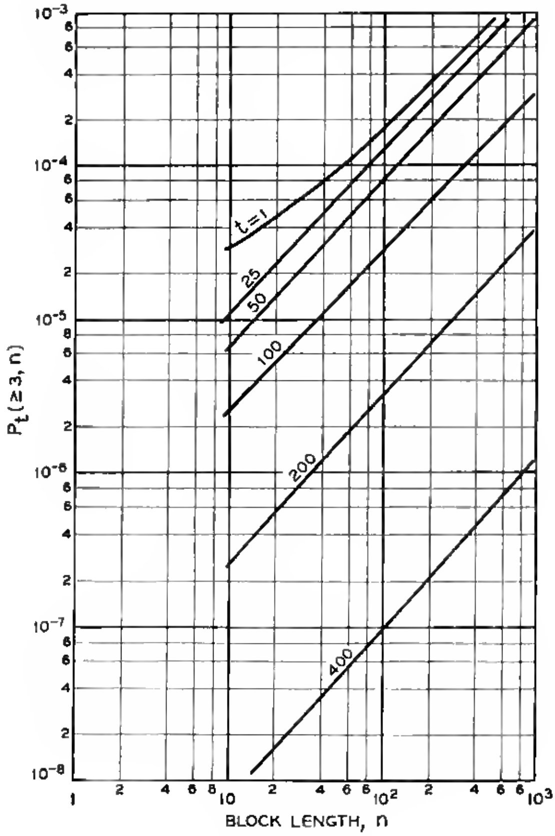

Fig. 14 — $P_t(\geqq 3,n)$ vs n.

exceed $P_t(\geqq 3,n)$, where n is the code's block length (and the interleaving constant is t).

Noting Fig. 13 and considering the Bose-Chaudhuri (31, 21) code,[10] which corrects all double errors, we conclude that interleaving with $t = 300$ would provide an error rate P_e of 10^{-7} or less. Such error rates are usually acceptable. The storage capacity required to obtain $t = 300$ for this block length is $31 \times 300 = 9{,}300$ bits.

VII. CONCLUSIONS

Representation of the telephone network by a combination of renewal-type channels may be accomplished by the specification of

slightly more than a dozen parameters. This has the advantages of being compact and convenient, and of admitting to extrapolation to large block lengths and giving accuracies which are more than adequate for most error-control evaluations.

Interleaving on the telephone network can significantly enhance the error-control effectiveness of error-correcting codes when the separation between adjacent bits of a code word is on the order of several hundred bits. The storage requirements for achieving such interleaving are excessive for general application, but, where computers or other special storage facilities are available, interleaving can provide an interesting trade-off between decoding complexity and simple storage.

REFERENCES

1. Gilbert, E. N., Capacity of a Burst-Noise Channel, B.S.T.J., **39**, Sept., 1960, p. 1253.
2. Berger, J. M., and Mandelbrot, B., A New Model for Error Clustering in Telephone Circuits, IBM J. Res. and Dev., **7**, July, 1963, p. 224.
3. Sussman, S. M., Analysis of the Pareto Model for Error Statistics on Telephone Circuits, IEEE Trans. Comm. Systems, June, 1963, p. 213.
4. Alexander, A. A., Gryb, R. M., and Nast, D. W. Capabilities of the Telephone Network for Data Transmission, B.S.T.J., **39**, May, 1960, p. 431.
5. Elliott, E. O., Estimates of Error Rates for Codes on Burst-Noise Channels, B.S.T.J., **42**, Sept., 1963, p. 1977.
6. Townsend, R. L., and Watts, R. N., Effectiveness of Error Control in Data Communications Over the Switched Telephone Network, B.S.T.J., **43**, Nov., 1964, p. 2611.
7. Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. I, 2nd ed., John Wiley and Sons, New York, 1959.
8. Morris, R., Further Analysis of Errors Reported in "Capabilities of the Telephone Network for Data Transmission," B.S.T.J., **41**, July, 1962, p. 1399.
9. MacWilliams, Jessie, Permutation Decoding of Systematic Codes, B.S.T.J., **43**, Jan., 1964, p. 485.
10. Peterson, W. W., *Error-Correcting Codes*, John Wiley and Sons, New York, 1961, p. 166.